

Quantum Monte Carlo Methods in Statistical Mechanics

Vilen Melik-Alaverdian and M.P. Nightingale

Department of Physics, University of Rhode Island, Kingston, Rhode Island 02881, USA

(October 14, 2000)

Abstract

This paper deals with the optimization of trial states for the computation of dominant eigenvalues of operators and very large matrices. In addition to preliminary results for the energy spectrum of van der Waals clusters, we review results of the application of this method to the computation of relaxation times of independent relaxation modes at the Ising critical point in two dimensions.

I. INTRODUCTION

The computation of eigenvalues and eigenstates of operators and large matrices is a ubiquitous problem. In this paper we review recent applications of the Quantum Monte Carlo methods that we have developed for this purpose. The reader is referred to other papers for introductory or more technical discussions of earlier work. [1–4]

II. MATHEMATICAL PRELIMINARIES

For an operator G , the power method can be used to compute the dominant eigenstate and eigenvalue, $|\psi_0\rangle$ and λ_0 .

This well-know procedure can be summarized as follows:

1. Choose a generic initial state $|u^{(0)}\rangle$ of the appropriate symmetry.
2. Iterate:

$$|u^{(t+1)}\rangle = \frac{1}{c_{t+1}} G |u^{(t)}\rangle, \quad (1)$$

where c_t puts $|u^{(t)}\rangle$ in standard form.

For projection time $t \rightarrow \infty$ the following is almost always true:

1. Eigenstate:

$$|u^{(t+1)}\rangle \rightarrow |\psi_0\rangle \quad (2)$$

2. Eigenvalue:

$$c_t \rightarrow \lambda_0 \quad (3)$$

To see this, expand the initial state in normalized eigenstates

$$|u^{(0)}\rangle = \sum_k w_k^{(0)} |\psi_k\rangle \quad (4)$$

with spectral weights $w_k^{(0)}$. Then $|u^{(t)}\rangle$ has spectral weights

$$w_k^{(t)} \sim \left(\frac{\lambda_k}{\lambda_0} \right)^t. \quad (5)$$

This method can be implemented by means of a Monte Carlo method and, unlike variational Monte Carlo, it has the advantage of producing unbiased results for large projection times t . The disadvantage is, however, that at the same time the statistical noise increases exponentially, unless G is a Markov (stochastic) matrix, or can be explicitly transformed to one. The statistical errors grow with the extent to which G fails to conserve probability, and to alleviate this problem, approximate dominant eigenstates can be used.

In the case of Markov matrices, computation of the dominant eigenvalue is of no interest, since it is equal to unity, but sampling the corresponding eigenstate has numerous applications.

A. Subspace iteration

Given a set of basis states, one can construct trial states as linear combinations to obtain approximate excited or, more generally, sub-dominant states and the corresponding eigenvalues. These are computed by solving a linear variational problem. In a Monte Carlo context, the Metropolis method can be used to evaluate the required matrix elements. Subsequently, a variation of the power method can again be used to remove systematically the variational bias. [5–7] Again, the price to be paid for reduction of the variational bias is increased statistical noise, a problem that can be mitigated by the use of *optimized* trial states.

The linear variational problem to be solved for the computation of excited states is the following one. Given n basis functions $|u_i\rangle$, find the $n \times n$ matrix of coefficients $D_i^{(j)}$ such that

$$|\tilde{\psi}_j\rangle = \sum_{i=1}^n D_i^{(j)} |u_i\rangle \quad (6)$$

are the “best” variational approximations for the n lowest eigenstates $|\psi_i\rangle$ of some Hamiltonian \mathcal{H} . In this problem we shall, at least initially, use the language of quantum mechanical systems, where one has to distinguish the Hamiltonian from the imaginary-time evolution operator $G = \exp(-\tau\mathcal{H})$. In the statistical mechanical application discussed below, we shall encounter only the equivalent of the latter, which is the Markov matrix governing the stochastic dynamics. In the expressions to be derived below, the substitution $\mathcal{H}G^p \rightarrow G^{p+1}$

will produce the expressions required for the statistical mechanical application, at least if we assume that the non-symmetric Markov matrix that appear in that context has been symmetrized, which can always be accomplished if detailed balance is satisfied.

Given these basis states, one seeks the “best” solution to the linear variational problem in Eq. (6) in the sense that for all i the Rayleigh quotient $\langle \tilde{\psi}_i | \mathcal{H} | \tilde{\psi}_i \rangle / \langle \tilde{\psi}_i | \tilde{\psi}_i \rangle$ is stationary with respect to variation of the coefficients of the matrix D . The solution is that the matrix of coefficients $D_i^{(j)}$ has to satisfy the following generalized eigenvalue equation

$$\sum_{i=1}^n H_{ki} D_i^{(j)} = \tilde{E}_j \sum_{i=1}^n N_{ki} D_i^{(j)}, \quad (7)$$

where

$$H_{ki} = \langle u_k | \mathcal{H} | u_i \rangle, \text{ and } N_{ki} = \langle u_k | u_i \rangle. \quad (8)$$

We note a number of important properties of this scheme. Firstly, the basis states $|u_i\rangle$ in general are not orthonormal. Secondly, it is clear that any nonsingular linear combination of the basis vectors will produce precisely the same results, obtained from the correspondingly transformed version of Eq. (7). The final comment is that the variational eigenvalues bound the exact eigenvalues from above, *i.e.*, $\tilde{E}_i \geq E_i$, where we assume $E_1 \leq E_2 \leq \dots$. One recovers *exact* eigenvalues E_i and the corresponding eigenstates, if the $|u_i\rangle$ span the same space as the exact eigenstates, or in other words, have no admixtures of more than n states.

The required matrix elements can be computed using the standard variational Monte Carlo method. The power method can subsequently be used to reduce the variational bias. Formally, one simply defines new basis states

$$|u_i^{(p)}\rangle = G^p |u_i\rangle \quad (9)$$

and substitutes these new basis states for the original ones. In quantum mechanical applications, where $G = \exp(-\tau\mathcal{H})$, the corresponding matrices

$$H_{ki}^{(p)} = \langle u_k^{(p)} | \mathcal{H} | u_i^{(p)} \rangle \quad (10)$$

and

$$N_{ki}^{(p)} = \langle u_k^{(p)} | u_i^{(p)} \rangle \quad (11)$$

can be computed by pure-diffusion Monte Carlo. [8] We note that, Monte Carlo yields these matrix elements up to an irrelevant overall normalization constant.

As an explicit example illustrating the nature of the Monte Carlo time-averages that one has to evaluate in this approach, we write down the expression for $N_{ij}^{(p)}$ as used for the computation of eigenvalues of the Markov matrix relevant to the problem of critical slowing down, discussed in detail in the next section. One estimates this matrix as

$$N_{ij}^{(p)} \propto \sum_t \frac{u_i(\mathbf{S}_t)}{\psi_B(\mathbf{S}_t)} \frac{u_j(\mathbf{S}_{t+p})}{\psi_B(\mathbf{S}_{t+p})}, \quad (12)$$

where the \mathbf{S}_t are configurations forming a time series that is designed to sample the distribution of a system in thermodynamic equilibrium, *i.e.*, the Boltzmann distribution ψ_B^2 . It

turns out that in this particular case, this distribution, the dominant eigenstate, has sufficient overlap with the magnitude of the sub-dominant states so that one can compute all matrix elements $N_{ij}^{(p)}$ simultaneously without introducing a separate guiding function [5].

The expression given in Eq. (12) yields the u/ψ_B -auto-correlation function at lag p . The expression for $H_{ij}^{(p)}$ is similar, and represents a cross-correlation function involving the configurational eigenvalues of the Markov matrix in the various basis states. Compared to the expressions one usually encounters in applications to quantum mechanical problems, Eq. (12) takes a particularly simple form in which products of fluctuating weights are absent, because one is dealing with a probability conserving evolution operator from the outset in this particular problem.

III. UNIVERSAL AMPLITUDE RATIOS IN CRITICAL DYNAMICS

Before continuing our general discussion, we temporarily change the topic to introduce stochastic dynamics of critical systems. What make such systems interesting, is that one can distinguish universality classes in which behavior does not depend on many of the microscopic details. For static critical phenomena, it is known that universality classes can be identified by dimensionality, symmetry of the order parameter, and the range of the interactions. For dynamical phenomena, there are additional features such as whether or not the dynamics is local or subject to conservation laws.

On approach of a critical point, the correlation length ξ diverges. The dynamical exponent z governs the corresponding divergence of the correlation time τ by means of the relation $\tau \propto \xi^z$. Since the critical exponent z is one of the universal quantities, it has been used to identify universality classes. Unfortunately, z does not vary by much from one universality class to another, and this poses a serious computational problem in terms of the accuracy required to obtain significant differences. One of the outcomes of the work reviewed here is that there are other quantities within computational reach, namely universal amplitude ratios. [4] These ratios may serve as additional, and possibly more sensitive identifiers of universality classes. We shall consider various systems belonging to a single universality class, and we assume that the representatives of the class are parameterized by κ .

If a thermodynamic system is perturbed out of equilibrium, different thermodynamic quantities relax back at a different rates. More generally, there are infinitely many independent relaxation modes for a system in the thermodynamic limit. The Monte Carlo methods reviewed here have been used to compute relaxation times of Ising models on square $L \times L$ lattices at the critical point. [4]

Let us denote by $\tau_{\kappa i}(L)$ the relaxation time of mode i of a system of linear dimension L . As indeed scaling theory suggests, it turns out that the relaxation time has the following factorization property

$$\tau_{\kappa i}(L) \approx m_{\kappa} A_i L^z, \quad (13)$$

where m_{κ} is a *non-universal* metric factor, which differs for different representatives of the same universality class as indicated; A_i is a *universal* amplitude which depends on the mode i ; and z is the *universal* dynamical exponent introduced above.

Formulated as a computational problem, one has the following. Suppose $\mathbf{S} = (s_1, \dots, s_{L^2})$, with $s_i = \pm 1$, is a spin configuration and $\rho_t(\mathbf{S})$ is the probability of finding \mathbf{S} at time t . The probability distribution evolves in time according to

$$\rho_{t+1}(\mathbf{S}) = \sum_{\mathbf{S}'} P(\mathbf{S}|\mathbf{S}') \rho_t(\mathbf{S}'). \quad (14)$$

The detailed structure of the Markov matrix P is of no immediate importance for the current discussion. All that matters is that it satisfies detailed balance, has the Boltzmann distribution ψ_B^2 as its stationary state. Also, P is a single-spin flip matrix, *i.e.* $P(\mathbf{S}|\mathbf{S}')$ vanishes if \mathbf{S} and \mathbf{S}' differ by more than a single spin. The desired relaxation time of mode i is given by

$$\tau_i(L) = -L^{-2} / \ln \lambda_i(L), \quad (15)$$

where λ_i is an eigenvalue of Markov matrix P . We obtained the previous expression by assuming a single-spin flip Markov matrix, so that the L^2 in the denominator produces a relaxation time measured in units of sweeps, *i.e.* flips per spin.

IV. TRIAL STATE OPTIMIZATION

To verify Eq. (13), it is important to obtain estimates that are exact within the range of the estimated error. For this purpose we use a set of optimized variational basis functions, to which we subsequently apply the projection procedure described in Section 2 to remove the variational bias.

As mentioned, the Monte Carlo projection increases the statistical noise, and the solution to this problem is to improve the variational basis functions. We shall now discuss how this is done and we consider the problem using the language of the Schrödinger equation.

We first consider the ground state and review how one can optimize a many-, say 50-parameter trial function $\psi_T(R)$. [9] The *local energy* $\mathcal{E}(R)$ is defined by

$$\mathcal{H}\psi_T(R) \equiv \mathcal{E}(R)\psi_T(R). \quad (16)$$

The variance of the local energy is given by

$$\chi^2 = \langle (\mathcal{H} - \bar{\mathcal{E}})^2 \rangle = \int |\psi_T(R)|^2 [\mathcal{E}(R) - \bar{\mathcal{E}}]^2 dR / \int |\psi_T(R)|^2 dR. \quad (17)$$

A property that we shall exploit later is that $\chi^2 = 0$ for any eigenstate, not just the ground state.

The following sums up the Monte Carlo optimization procedure for a single state:

1. Sample R_1, \dots, R_s from ψ_T^2 a typical sample size has $s \approx 3,000$.
2. Approximate the integrals in Eq. (17) by Monte Carlo sums.
3. Minimize χ^2 as follows, while keeping this sample *fixed*. For each member of the sample R_1, \dots, R_s :

4. Compute $\psi_T(R_1), \dots, \psi_T(R_s)$.
5. Compute $\mathcal{H}\psi_T(R_1), \dots, \mathcal{H}\psi_T(R_s)$.
6. Find $\bar{\mathcal{E}}$ from least-squares fit of

$$\mathcal{H}\psi_T(R_\sigma) = \bar{\mathcal{E}}\psi_T(R_\sigma), \sigma = 1, \dots, s. \quad (18)$$

7. Minimize the sum of squared residues of Eq. 18.¹

This procedure can be generalized immediately to a *set* of basis functions, as required to implement Eq. (6). The only new ingredient is a guiding function ψ_g^2 that has sufficient overlap with all basis states used in the computation. For this purpose one can conveniently use the groundstate raised to some appropriate power less than unity.

This yields the following algorithm to optimize basis states for n dominant eigenvalues:

1. Sample R_1, \dots, R_s from ψ_g^2 .
2. Compute the arrays

$$\begin{pmatrix} u^{(1)}(R_1) \\ u^{(2)}(R_1) \\ \vdots \end{pmatrix}, \dots, \begin{pmatrix} u^{(1)}(R_s) \\ u^{(2)}(R_s) \\ \vdots \end{pmatrix}. \quad (19)$$

3. Compute the arrays

$$\begin{pmatrix} \mathcal{H}u^{(1)}(R_1) \\ \mathcal{H}u^{(2)}(R_1) \\ \vdots \end{pmatrix}, \dots, \begin{pmatrix} \mathcal{H}u^{(1)}(R_s) \\ \mathcal{H}u^{(2)}(R_s) \\ \vdots \end{pmatrix}. \quad (20)$$

4. Find the matrix elements $\bar{\mathcal{E}}_{ij}$ from the appropriately weighted least-squares fit to

$$\mathcal{H}u^{(i)}(R_\sigma) = \sum_{j=1}^n \bar{\mathcal{E}}_{ij} u^{(j)}(R_\sigma), \sigma = 1, \dots, s. \quad (21)$$

5. Vary the parameters to optimize the fit, as explained below.

In case of a perfect fit, the eigenvalues of the truncated Hamiltonian matrix $\mathbf{E} = (\bar{\mathcal{E}}_{ij})_{i,j=1}^n$ are the required eigenvalues, but in real life one has to optimize the parameters of the basis functions, which can be done as follows:

1. Divide the sample in blocks and compute one Hamiltonian matrix \mathbf{E} per block.

¹Once the parameters are changed from the values they had in step 1, one should use an appropriately weighted sum of squared residues. [9]

2. Minimize the variance of the \mathbf{E} -spectra over the blocks.

The variance vanishes if the basis functions $u^{(i)}$ are *linear combinations* of n eigenstates of \mathcal{H} . This gives rise to a computational problem, *viz.*, the variance is near-invariant under linear transformation of the $u^{(i)}$. This approximate “gauge invariance” gives rise to near-singular, non-linear optimization problem. This can be avoided by simultaneously minimizing the variance of both the spectrum of the “local” Hamiltonian matrix \mathbf{E} and the local energy \mathcal{E} of the individual basis functions.

Finally, the variational bias of the eigenvalue estimates obtained with the optimized basis states is reduced by using Monte Carlo to make the substitution discussed previously

$$|u^{(i)}\rangle \rightarrow e^{-\mathcal{H}\tau} |u^{(i)}\rangle. \quad (22)$$

For this purpose, one has to use the short-time approximation of $\exp(-\mathcal{H}\tau)$. [5] To apply the preceding scheme to the problem of critical dynamics, all one has to do is to make use of the fact the analog of the quantum mechanical evolution is the symmetrized Markov \hat{P} of stochastic dynamics, which is defined as

$$\hat{P}(\mathbf{S}|\mathbf{S}') = \frac{1}{\psi_{\mathbf{B}}(\mathbf{S})} P(\mathbf{S}|\mathbf{S}') \psi_{\mathbf{B}}(\mathbf{S}'), \quad (23)$$

in terms of which we have the correspondence

$$e^{-\mathcal{H}\tau} \rightarrow \hat{P}^t. \quad (24)$$

V. XE TRIMER: A TEST CASE

As an example that illustrates the accuracy one can obtain by means of the optimization schemes discussed above, we present results for a Xe trimer interacting via a Lennard-Jones potential. To be precise, we write the Hamiltonian of this system in *reduced* units as

$$\mathcal{H} = -\frac{1}{2m} \nabla^2 + \sum_{i < j} (r_{ij}^{-6} - 2) r_{ij}^{-6}, \quad (25)$$

where the r_{ij} denote the dimensionless interparticle distances. We *define* Xe to correspond to $m^{-1} = 7.8508 \times 10^{-5}$, which probably to four significant figures [10] agrees with Leitner *et al.* [11]

Table I shows results for variational energies of the lowest five completely symmetric states of a Lennard-Jones Xe trimer. The results are compared with results obtained by the discrete variable representation truncation-diagonalization method. [11] The basis functions used in this computation are of the same general form used in earlier work with an additional polynomial prefactor for excited states. [12,13]

Clearly, we obtain consistently lower reduced energies, which we attribute to lack of convergence of the results of Leitner *et al.* [14]

TABLES

k	E_k	σ	Leitner <i>et al.</i>
0	-2.845 241 50	1×10^{-8}	-2.844
1	-2.724 955 8	1×10^{-7}	-2.723
2	-2.675 065	1×10^{-6}	-2.664
3	-2.608 612	2×10^{-6}	-2.604
4	-2.592 223	3×10^{-6}	-2.580

TABLE I. Variational reduced energies compared with estimates of Leitner *et al.*

VI. CRITICAL POINT DYNAMICS: RESULTS

Next we briefly address the issue of the choice of trial functions for the eigenstates of symmetrized Markov matrix \hat{P} . We write

$$u(\mathbf{S}) = f(\mathbf{S}) \times \psi_{\mathbf{B}}(\mathbf{S}). \quad (26)$$

For the modes we considered, $f(\mathbf{S})$ was chosen to be a rotationally and translationally invariant polynomial in long-wavelength Fourier components of \mathbf{S} , the lowest-order one of which is simply the magnetization. Corresponding to the order parameter and energy-like modes, we considered polynomials either odd or even under the transformation $\mathbf{S} \rightarrow -\mathbf{S}$.

We briefly discuss some of the results that illustrate the validity of Eq. (13). Figure 1. shows plots of the effective amplitudes for the three dominant odd, and two dominant even modes of three different Ising models on $L \times L$ lattices. Of the three Ising models we studied, the first one, the NN model, had nearest-neighbor couplings only. The other two also had next-nearest-neighbor couplings. In one of them, the equivalent neighbor or EQN model, both couplings were of equal ferromagnetic strengths. In the third or NEQ model, the nearest-neighbor coupling was chosen ferromagnetic and of twice the magnitude of the antiferromagnetic next-nearest-neighbor coupling.

FIGURES

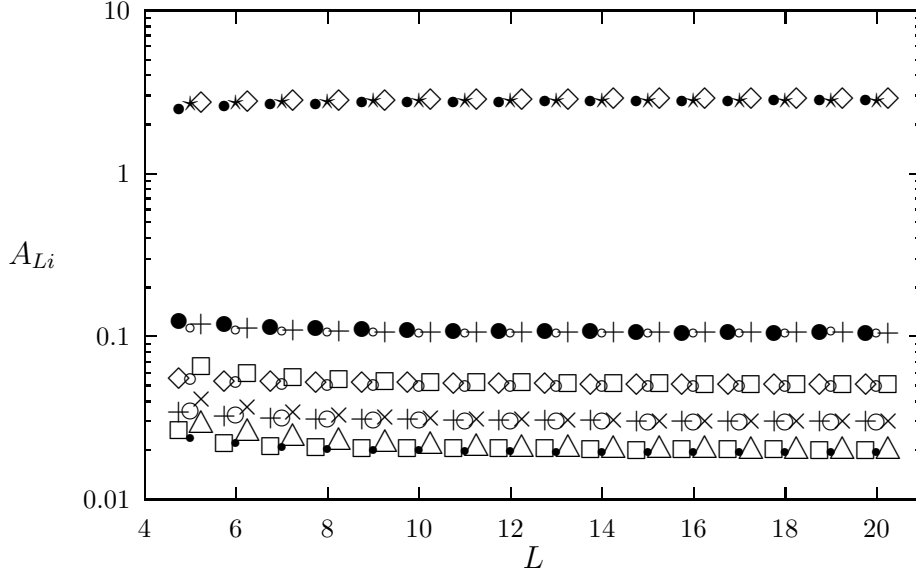


FIG. 1. Universality of relaxation-time amplitudes, shown in a plot of the effective, size-dependent amplitudes A_{Li} on a logarithmic scale. To separate data points for the three models, the NEQ data were displaced to the left and the EQN data to the right. The data collapse predicted by Eq. (13) was produced by fitting the metric factors of the NN and NEQ models. Amplitudes of odd and even states alternate in magnitude.

To obtain estimates of the amplitudes of the relaxation modes, we fit the computed correlation times to expressions of the form

$$\tau_i(L) \approx L^z \sum_{k=0}^{n_c} \alpha_{ki} L^{-2k}. \quad (27)$$

In our computation of the non-universal metric factors, this quantity was set equal to unity by definition for the EQN model. Table II shows the metric factors computed for each mode separately as the ratio of the computed amplitudes. In agreement with the scaling prediction in Eq. (13), the computed metric factors depend only on the model but not on the mode.

TABLE II. Non-universal metric factors m_κ , as defined in Eq. (13), computed for the NN and NEQ models. The modes indicated by o1, o2, and o3 are odd under spin inversion; the remaining two, e2 and e3, are even.

	NEQ	NN
o1	2.389(1)	1.5569 (5)
e2	2.394(2)	1.5569 (5)
o2	2.393(2)	1.5567 (6)
e3	2.391(2)	1.554 (2)
o3	2.385(4)	1.554 (2)

Finally we mention that the spectral gaps of the Markov matrix vary over a considerable range

$$1 - \lambda_i(L) \approx L^{-(d+z)} \approx L^{-4.17}, \quad (28)$$

i.e. from approximately 3×10^{-3} for $L = 4$ to 3×10^{-6} for $L = 21$. For details of the numerical analysis based on Eq. (27) we refer the interested reader to Ref. [4]. Suffice it to mention that the value obtained for the universal dynamic critical exponent z featured in Eq. (13) is $z = 2.167 \pm 0.002$ which is indistinguishable from $13/6$.

ACKNOWLEDGMENTS

This work was supported by the (US) National Science Foundation through Grants DMR-9725080 and CHE-9625498. It is a pleasure to acknowledge helpful e-mail exchanges with David Leitner.

REFERENCES

- [1] M.P. Nightingale and C.J. Umrigar, *Monte Carlo Eigenvalue Methods in Quantum Mechanics and Statistical Mechanics*, in Advances in Chemical Physics, Vol. 105, Monte Carlo Methods in Chemistry, edited by David M. Ferguson, J. Ilja Siepmann, and Donald G. Truhlar, series editors I. Prigogine and Stuart A. Rice, Chapter 4 (John Wiley and Sons, New York 1999).
- [2] M.P. Nightingale and C.J. Umrigar (eds.), *Quantum Monte Carlo methods in Physics and Chemistry*, NATO Science Series, Series C: Mathematical and Physical Sciences - Vol. 525 (Kluwer Academic Publishers, Dordrecht, 1998).
- [3] M. P. Nightingale and H.W.J. Blöte, *Phys. Rev. Lett.* **76**, 4548 (1996).
- [4] M. P. Nightingale and H.W.J. Blöte, *Phys. Rev. Lett.* **80**, 1007 (1998). Also see <http://xxx.lanl.gov/abs/cond-mat/9708063>.
- [5] D.M. Ceperley and B. Bernu, *J. Chem. Phys.* **89**, 6316 (1988).
- [6] B. Bernu, D.M. Ceperley, and W.A. Lester, Jr., *J. Chem. Phys.* **93**, 552 (1990).
- [7] W.R. Brown, W.A. Glauser, and W.A. Lester, Jr., *J. Chem. Phys.* **103**, 9721 (1995).
- [8] M. Caffarel and P. Claverie, *J. Chem. Phys.* **88**, 1088 (1988); *ibid.* p. 1100. Also see S. Baroni and S. Moroni in Ref. [2].
- [9] C.J. Umrigar, K.G. Wilson, and J.W. Wilkins, *Phys. Rev. Lett.* **60**, 1719 (1988); *Computer Simulation Studies in Condensed Matter Physics*, edited by D.P. Landau, K.K. Mon, and H.-B. Schüttler, Springer Proceedings in Physics Vol. 33 (Springer-Verlag, Berlin, 1988), p.185.
- [10] This number was obtained by using mass and Lennard-Jones parameters of Ref. [11], as given in Table I and the fundamental physical constants of E.R. Cohen and B. N. Taylor, in the supplement to the August 1999 issue of Physics Today, BG5.
- [11] D.M. Leitner, J.D. Doll, and R.M. Whitnell, *J. Chem. Phys.* **94**, 6644 (1991).
- [12] Andrei Mushinski and M. P. Nightingale, *J. Chem. Phys.* **101**, 8831 (1994).
- [13] M. Meierovich, A. Mushinski, and M.P. Nightingale, *J. Chem. Phys.* **105**, 6498 (1996).
- [14] Indeed recent, improved computations by Leitner (private communication) produce lower energies than the results quoted in Table II of Ref. [11].